

Open Research Online

The Open University's repository of research publications and other research outputs

Live Blackboxes: Requirements for Tracking and Verifying Aircraft in Motion

Conference or Workshop Item

How to cite:

Yu, Yijun; Yang, Mu and Nuseibeh, Bashar (2017). Live Blackboxes: Requirements for Tracking and Verifying Aircraft in Motion. In: SCiA 2017 : 4th Software Challenges in Aerospace Symposium, 9-13 Jan 2017, Grapevine, Texas.

For guidance on citations see [FAQs](#).

© [not recorded]



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Accepted Manuscript

Link(s) to article on publisher's website:
<http://dx.doi.org/doi:10.2514/6.2017-0884>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

Live Blackboxes: Requirements for Tracking and Verifying Aircraft in Motion

Yijun Yu¹ Mu Yang¹ Bashar Nuseibeh^{1,2}

¹*Department of Computing and Communications, The Open University, UK*

²*Lero the Irish Software Research Centre, Limerick, Ireland*

The Malaysian Airlines (MH370) aircraft went missing somewhere over the Indian Ocean two years ago. After intensive search since then, international team still has not been able to locate any first-hand evidence from the missing plane’s flight data recorders (also known as ‘blackboxes’). To mitigate similar problems, a proposal has been made to analyse live streamed flight data using cloud computing; however, satellite communication is constrained by bandwidth and scalability challenges.¹ In this paper, we propose five requirements for addressing these challenges. These requirements frame a class of monitoring problems that share some similar accuracy concerns around safety and security. We evaluate these requirements to assess the readiness of the proposed technology - which we call “live blackboxes” – by using actual global scale data and performing an analysis of different live streaming intervals. Preprocessing with a locality-sensitive hashing function, it results in reduction of the required bandwidth by 4.75 times. Therefore, to track and verify all civilian aircraft in motion, the scalability requirement could be satisfied by satellite communications. While the paper focuses on a particular problem in air traffic management, we speculate similar requirements for the continuous monitoring of critical systems.

I. Introduction

The last decade of aviation saw improving safety and security, until 2014, when 900 people lost their lives or remained missing (see Figure 1).² The root causes for all the crashes except for MH370³ are already

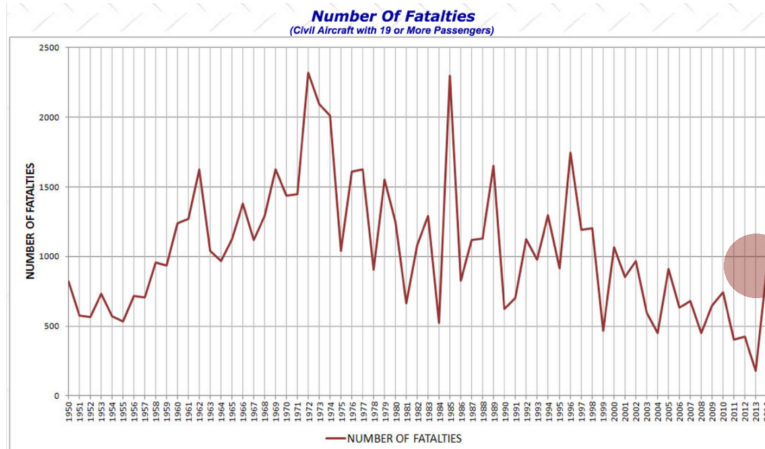


Figure 1: Number of Casualties spiked up in 2014

known because Flight Data Recorders (FDRs), also known as the *blackboxes*, were found eventually at the crash sites. Essential to post-mortem forensic diagnosis, such boxes contain well-protected information about the cockpit communications amongst pilots and with ground controllers, and a number of operating parameters, e.g., longitude, latitude, altitude, vertical and horizontal velocities. As a standard practice, once the causes are confirmed by diagnosing the data extracted from the blackboxes, aviation industry could

mitigate similar failures by releasing concrete guidance necessary to the safety of future flights. Delays and resources in recovering the blackboxes are often obstacles to timely investigation of the crashes, including e.g., the recent EgyptAir Flight 804.⁴

However, it is not yet known what has happened to the missing MH370. Its last hourly ping signals to an Inmarsat satellite suggested that the aircraft actually headed towards the southern Indian Ocean, according to the Doppler effect of drifting frequency caused by high-speed movements of the signal source. Though the international team shifted the search and rescue effort immediately from the South China Sea towards the Indian ocean, it was hard to locate the blackboxes under the ocean bed because the batteries for signalling their location lasted no longer than 30 days. As the batteries ran out, a large area was still unexplored. After spending tremendous amount of resources, a piece of the wings was found on 29 July 2015 at Reunion Island, which floated there from a potential crash site driven by the ocean currents.^{5,6} A more recent discovery of a Boeing 777 part ashore Mozambique on March 2, 2016,⁷ which may contain some more clues. However, without locating the blackboxes that recorded critical flight data, the search has to continue.

Invented in the 1950s, blackboxes are still being used widely as the primary source of forensic investigation for aviation safety and security. In the wake of Air France Flight 447 incident earlier in 2009, Clive Irving proposed to consider replacing blackboxes with some new technology.⁸ Due to the missing blackboxes in four incidents between 1983 and 2010, van’t Kaar et al⁹ have proposed to investigate the legal aspect of online blackboxes; however, these suggestions had not been put into practice. Since the MH370 incident, the International Telecommunication Union (ITU) has adopted a *live streaming* proposal¹⁰ in its guidance for future aviation communications.¹¹ With widely used remote communication capabilities, even smart-phone owners can already detect the location of their phones and wipe data out if they are lost. It is reasonable to expect that engineers do the same for aircraft in motion by advancing the technology.

Tracking one smart-phone seems to be easy but it is not. Millions of smart-phone owners, distributed globally, may need this capability simultaneously. To be able to fulfil this simple requirement it requires vendors to invest heavily in *cloud computing*, the on-demand technology to use large amount of computation resources when needed. Additionally, for aircraft flying over the oceans, radar signals are not available and satellite communications seem to be the most reliable ways to relay location information. However, there is already a bandwidth problem concerning the need of tracking all aircraft including those designated journeys on land: e.g., the MH370 aircraft was destined from Kuala Lumpur to Beijing, rather than above the Indian Ocean. Here, we propose five “live blackbox” requirements, namely, tracking, prediction, verification, scalability, and liveness, to enhance aviation safety and security.

The research challenge on current technology is to achieve all these requirements together. We attempt to show their feasibility, using continuously captured time-series of radar data collected from 21,616 globally operating aircraft, in a duration of 3 weeks between January 27 2016 and February 17 2016. This amounts to 69,137,191 minute-by-minute update locations. Our main contributions can be summarised as follows:

- Highlighting the importance of tracking requirements for aviation safety and security of individual aircraft, in addition to collision-avoidance between two aircraft;
- Estimating the amount of satellite bandwidth required for communicating the locations by collecting time-series data from flight radars;
- Predicting the precise location on the next interval based on either previous locations on the same trajectory or the previous trajectories;
- Using locality-sensitive hashing (LSH) technique to verify the predicted trajectory by accuracy rather than by precision against anomalies, requiring smaller bandwidth; and
- Assessing the reduction of necessary bandwidth to address the scalability requirement of satellite communications.

For safety, all systems are verified to detect anomalies to the projected trajectory; for security, locality-sensitive instead of crypto hashing is used for continuous phenomena in physical domain; and for accuracy, the boundary check decision is no longer based on unnecessarily high precision, in order to accommodate high demand on sampling frequency (i.e., liveness) and scalability.

The remainder of the paper is organised as follows. Section 2 explains why the state-of-the-art satellite communications have not been deployed for tracking aircraft in motion; Section 3 presents the definitions of five live blackbox requirements to enable time-series analysis for predicting trajectories and detecting anomalies responsively for a large volume of aircraft. Section 4 reports our pilot experiments to evaluate the feasibility of the proposed live blackbox requirements using flight radar data and locality-sensitive hashing. It also discusses threats to validity in those experiments. Section 5 presents related work ranging from

monitoring problems, security requirements, to existing air traffic control-specific requirements. Finally, Section 6 concludes the study and indicates a few directions of further research.

II. Tracking Aircraft in Motion: Data in Communications

To track aircraft in motion, the basic requirement is to get sufficiently high quality flight data.

Now there are two fundamental ways in collecting flight data, via satellites and radars. Similar to how mobile phones locate themselves by aligning with Global Position Systems (GPS) satellites, aircraft GPS locations are obtained. Locating GPS is anonymous and does not require additional satellite communication bandwidth. On top of the GPS location data (i.e. longitude, latitude, altitude), modern aircraft also broadcast their identities and the identities of the associated air-traffic control radar zone, flight speed, vertical speed using the Automatic Dependent Surveillance-Broadcast (ADS-B) radar technology since 2013. Companies such as FlightRadar^a provide low cost wireless ADS-B receivers for clients to gather radar signals of the aircraft in range. Through crowd-sourcing owners of such systems upload the received data to FlightRadar servers, which lead to a live updated status of global aircraft. When ADS-B signals are not available, FlightRadar uses the Multi-LATeration (MLAT) technique to aggregate multiple observations into aircraft GPS locations.

In this work, we acquire flight tracking data using the FlightRadar API. The data obtained are minute-by-minute, and they cover populated area in the globe, except for scarcely populated areas such as forests, desserts and oceans. In the rest of the paper, we will base our analysis on what is available from the data collection. To capture efficiently all live aircraft location data, we also increased the intervals from 1 minute to 5 minutes such that the experimental server could handle all the information continuously.

For each data point, we get longitude, latitude, altitude, horizontal speed, vertical speed, as well as flight number, origin airport, destination airport, radar zone, and last update time.

Each aircraft has variable numbers of data-points during the three weeks period, on average there are 3,198 data points per aircraft. An aircraft, identified by “06A0BC”, is chosen for the most data points amongst the 21,616 aircraft over three weeks. The aircraft has 25,714 data points. Its latitude, longitude, and altitude changes are shown in Figure 2.

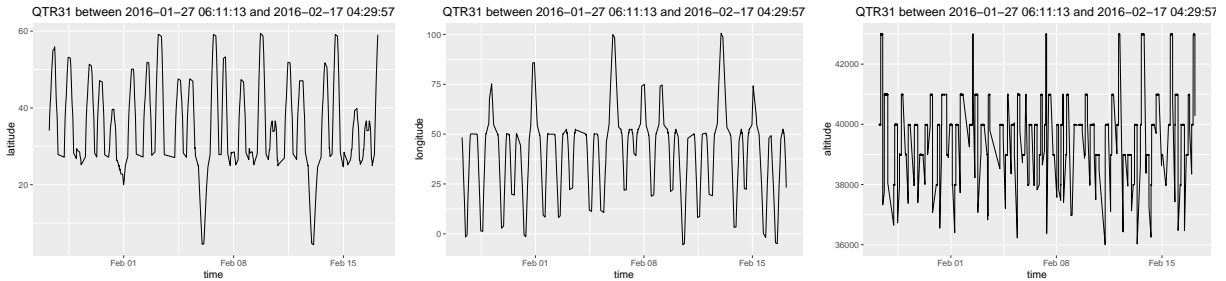


Figure 2: Live changes of the latitude of aircraft 06A0BC over 3 weeks

From the flight id of aircraft 06A0BC^b, it is clear that the aircraft is operated by Qatar Airlines on multiple (origin, destination) pairs of journeys. The Boeing 787 flew from Doha, Qatar to Edinburgh, London in the UK, Stockholm in Sweden, Zurich in Switzerland, Kuala Lumpur in Malaysia, etc., and back forth. That is why there are segments on the curves with different latitude/longitude end-points. Observable from the curves of longitude/latitude is the fluctuations at the phase of landing, where the flights often hover around the airport to wait for an opportunity to touch down on the runway. Worth noting are the sudden changes of altitude during these flights. From the non-zero speed and altitude data, it is also observable that the flight data obtained are mostly in-flight, while the ADS-B broadcasts are turned on. In fact, the altitude in cruise mode are fluctuating, ranging between 36,000 and 44,000 feet.

^a<https://www.flightradar24.com/>

^bCan be monitored live at <https://planefinder.net/flight/z.NO-06A0BC>

III. Requirements for Live Blackboxes

With the data collected from Flight Radars, it is possible to articulate the requirements for live blackboxes more precisely, which would help with understanding the pilot experiments we report next. Informally, Figure 3 illustrates these requirements in the context of one flight journey of an aircraft. The tracking requirement shall record the flight data and communicate to a radar or a satellite; the prediction requirement shall tell the motion of the aircraft based on the previous data points in the trajectory; the verification requirement shall be able to determine whether the flight is on track; the liveness requirement shall report any anomaly within a time interval; and finally, scalability requirement shall guarantee the bandwidth of satellite communications is sufficient for fulfilling the above requirements.

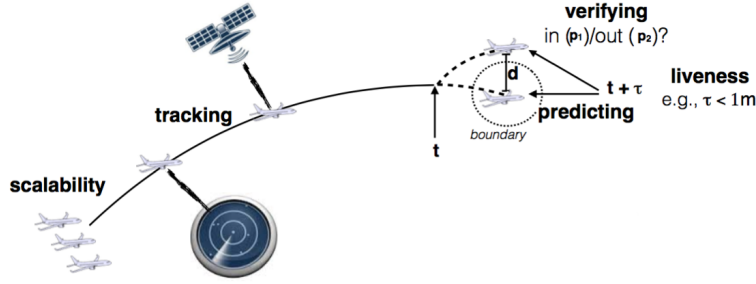


Figure 3: Illustration of the live blackbox requirements

These requirements are defined by the data points collected, which lead to the overall requirement for live blackboxes. The first requirement is to track an aircraft in motion.

Requirement 1 Tracking. *An aircraft is tracked on an interval of τ if from every time stamp t during its journey, there is at least one data point(s) between $(t, t + \tau)$. The data point needs to contain parameters to uniquely identify the flying object on the physical space at the time-point.*

Our radar dataset could satisfy this requirement to the interval of $\tau = 1$ minute. In the wake of the MH370 incident, U.N. aviation agency has approved now for International Civil Aviation Organisation (ICAO) to propose a new standard that aircraft on oceans shall update their locations every 15 minutes,¹² rather than every 1 hour. In other words, $\tau=15$ minutes. Generally speaking, the smaller the interval is, the more precise it would help track the aircraft.

Requirement 2 Prediction. *Given the location data of an aircraft up to time t , the location at the next interval $t + \tau$ could be predicted as $L'(t + \tau)$. Compared to the location in reality at $t + \tau$, the error of prediction e can be judged by a threshold of physical distance: when $\|L'(t + \tau) - L(t + \tau)\| < e$, the prediction is accurate; otherwise it is faulty. Given the threshold e , the overall prediction requirement aims to find a model to lower the chance of faulty predictions.*

There are many ways to implement this prediction requirement, such as regression models and decomposition analysis, given the datasets we collected.

It is our hypothesis that when the tracking interval τ is smaller, the prediction model is more likely to be correct with respect to a given threshold e . Part of our experiments will be to check this hypothesis by intentionally ignoring some data points of a smaller intervals to see whether the accuracy of prediction of the trajectory could be affected.

When intervals get smaller, however the number of data points on a given trajectory will definitely increase, hence the scale of telecommunications to transmit the location data to analyse. Once the data arrives at the ground, it is considered acceptable to use cloud computing to deal with the computational scalability issues. Hence, the main technical concern is how efficient telecommunication can scale up to include all the hundreds of thousands of aircraft.

Requirement 3 Scalability. *Given the average duration of a flight D , the amount of bandwidth required for communicating location data of n aircraft in motion at an interval of τ is estimated to be at the scale of $n \times D/\tau \times B$, where B is the number of bytes for representing the location data.*

For the experiments we will use at least 3 independent flight data variables (longitude, latitude, altitude), and consider each variable in double precision require $B = 7 + 7 + 5 = 19$ bytes. Using the 3 weeks data we collected from flight radars, at $\tau=1$ minute interval, there are at least $69,137,191 \times 19$ bytes = 1,313,606,629 bytes of communication spanning across 3 weeks of 30,240 minutes. The entire radar datasets, including aircraft ID, speed and timestamps, would otherwise costs more than 10GB after compression. On average, it is 43,439 bytes per minute.

Broadcasting at $\tau = 5$ minutes, it satisfies the ICAO satellite communication requirement, and the bandwidth requirement is reduced to 3,048 bytes per minute (Bpm), or 24,387 bits per second (bps) (one byte is 8 bits). Having been reviewed in the ICAO, apparently, this rate is acceptable for the current satellite communication technology and the cost can be shared by airlines. However, if we would decrease the τ interval below that of the current radar datasets, while still maintaining the capability in predicting the location of aircraft in motion, it requires some innovative solution.

One could use hashing function to reduce the bandwidth requirement because any change of the input value will almost certainly lead to a different hashing value. However, verifying objects in the physical domain cannot be that simple, because a perturbation of the input caused by small gust of wind, for example, should not be considered as catastrophe for the aircraft motion to be considered as abnormal. Instead, a locality-sensitive hashing (LSH) may provide a way to achieve the compression of inputs without losing the capability in verifying the locations. Inspired by the concepts of LSH,¹³ the definition of the verification requirement is presented below:

Requirement 4 Verification. *To tell whether an aircraft in motion is in a location far away from the predicted trajectory, the locality-sensitive verification is to tell, with $c \geq 1$ that (1) the probability is no smaller than P_1 for the aircraft to be within the distance d of where it is predicted, and (2) the probability is at most P_2 for the aircraft to be outside the distance cd , where $P_1 > P_2$.*

The two probability parameters P_1 and $1 - P_2$ could tell us respectively how confident one could be when the hashed location is known. If we choose $c = 1$ it will be crisp (i.e., binary) in verifying the location either matches or not matches with the predicted trajectory.

By definition, a family of LSH functions could satisfy the verification requirement above. However, we aim to find a function that requires smaller number of bytes for communication, without missing the opportunity of detecting abnormal motions. Once an LSH function is chosen, both on-board aircraft and the ground tracking system need to use the same function consistently in this protocol.

From the aircraft to satellite and to the receiving ground stations, telecommunication imposes delays. According to the vision of 5G standard,¹⁴ the next generation of telecommunication would introduce an adaptive layer to allow for the *interpolation* of data points in-between two consecutively transmitted points. Otherwise, the speed limit of light imposes a 50 ms limit for the receiver to get the data packets from the sender. In other words, if the interval τ is larger than 50 ms, it is still possible to fill it up with projected values in the predicted trajectory. The aim of such interpolations is to achieve the liveness requirement.

Requirement 5 Liveness. *Given the data points collected from the past, the live prediction is to tell, with as short delay τ as possible, the location of aircraft in motion.*

When $\tau < 50$ ms, it is required to apply interpolations. Although the current satellite communication may not yet support this level of transmissions, we are speculating that in the future it is indeed feasible. In order to estimate whether the liveness through interpolation could sacrifice predictability and verifiability requirements mentioned earlier, we also conducted some experiments using flight radar data, on larger pairs of $\tau=1$ minute and $\tau = 5$ minutes, just to see whether it is feasible to foresee such effect.

These five requirements are not independent. More specifically, the interval parameter τ occurs in the tracking and prediction requirements. A smaller τ could make the tracking requirement difficult to be satisfied while increasing the accuracy of prediction. In addition, the scalability requirement is determined by both the interval and the prediction accuracy (i.e. ϵ): the smaller the interval and error tolerance, the more data points are needed and hence more demanding to satisfy scalability. The verification requirement depends on the boundary and locality sensitive probability threshold c . Finally, the liveness prediction requirement relates to the prediction and scalability requirements. It is interesting to investigate whether better scalability (e.g., via larger interval and error threshold) would make it easier to satisfy the liveness requirement. In the pilot experiments, we aim to answer these questions using realistic data.

IV. Pilot Experiments

In the interest of understanding the proposed requirements, results of a pilot experiment are reported.

A. Predicting aircraft trajectory in time-series

In the first experiment we evaluate whether statistical methods are effective in predicting trajectories of aircraft in motion. To this end, we construct prediction models using collected flying data and cast the models as a regression problem and a seasonal decomposition problem, respectively.

Linear regression is the one of the most commonly used predictive analysis. Regression estimates are used to describe data and to explain the relationship between one dependent variable and at least one independent variables. Given a time-series data set $\{(y, x_1, \dots, x_p)_i | i = 1, \dots, t\}$ of t samples, a linear regression model assumes that the relationship between the dependent variable y and the p -vector of independent variables \mathbf{x} is linear. This relationship is modelled through the following equation.

$$y(t) = \beta_0 + \beta_1 x_1(t) + \beta_2 x_2(t) + \dots + \beta_p x_p(t) + \epsilon \quad (1)$$

where β_0, \dots, β_p are the regression coefficients, ϵ is the error term that captures all other factors which influence the dependent variable y other than the independent variables \mathbf{x} .

In the trajectory prediction scenario, we want to predict the longitude (or latitude/altitude) of an aircraft at time $t+\tau$ when there is no collected data. Taking longitude for example, the factors that influence longitude can include previous longitude, latitude, altitude, and the horizontal and vertical speeds of the aircraft which were collected at time intervals of τ . Therefore, we have the dependent variable y denoted by $longitude_{t+\tau}$ and the vector \mathbf{x} includes $p = 5$ independent variables: the previous longitude, latitude, altitude, horizontal speed and vertical speed. That is, $\mathbf{x} = \{longitude_t, latitude_t, altitude_t, speed_t, vspeed_t\}$. To predict location at each time stamp $t + \tau$ where $t = i\tau$ and $i \in \mathbb{Z}^+, i \geq 1$, we take the following data set as training set^c, whose size incrementally increases as i increases:

$$\{longitude_t, longitude_{t-\tau}, latitude_{t-\tau}, altitude_{t-\tau}, speed_{t-\tau}, vspeed_{t-\tau}\}_{t=2\tau}^{t=i\tau}. \quad (2)$$

Because the training set changes at each time stamp, the obtained regression coefficients and error terms are different at different timestamps, that is $\beta_0(t), \dots, \beta_p(t)$ and $\epsilon(t)$. We have applied the regression approach to the data of aircraft 06A0BC and predicted one of its trajectories flying from Doha to Edinburgh airport on 27th January 2016.

To see how the tracking requirement τ affects the accuracy of prediction, we manually removed some data points to vary the interval of data collection. Figure 4 shows the performance of the regression approach when the data collection interval τ is at every one and five minutes. The prediction for longitude and latitude is always good. However, altitude prediction is not accurate at a sudden increase or drop in the data, especially when the data collection interval τ is five minutes. The fine-grained data (i.e., collected at every one minute) improves the prediction because there are more data in the training data set.

We showed the prediction error when τ is 5 minutes in Figure 5 by computing the difference between the actual and predicted values. The error of predicting longitude and latitude decreases as time increases because the regression model has more data to learn the correct coefficients. However, the error for altitude is almost zero at the beginning until there is a sudden increase in altitude. This is because altitude does not change for a long time at the beginning and this can be quickly captured by the regression model.

An alternative to the regression approach is the seasonal decomposition of time series, which is widely used for forecasting time series when time series of previous periods are available.

The decomposition approach divides a time series into *notional* components,¹⁵ including seasonal, trend, cyclic and random (containing anything else in the time series). Assuming an additive model, a time series y_t can be written as follows.

$$y_t = S_t + T_t + C_t + R_t \quad (3)$$

where y_t is the time-series data at time t and S_t (resp. T_t , C_t and R_t) is the seasonal (resp. trend, cyclic and random) component at time t .

In the prediction scenario, we used the historical trajectories of the same air route of an aircraft as the time-series training data and predicted the future trajectories using the decomposition approach. More

^cThere is no previous data that can be trained at time $t = \tau$ where $i = 1$.

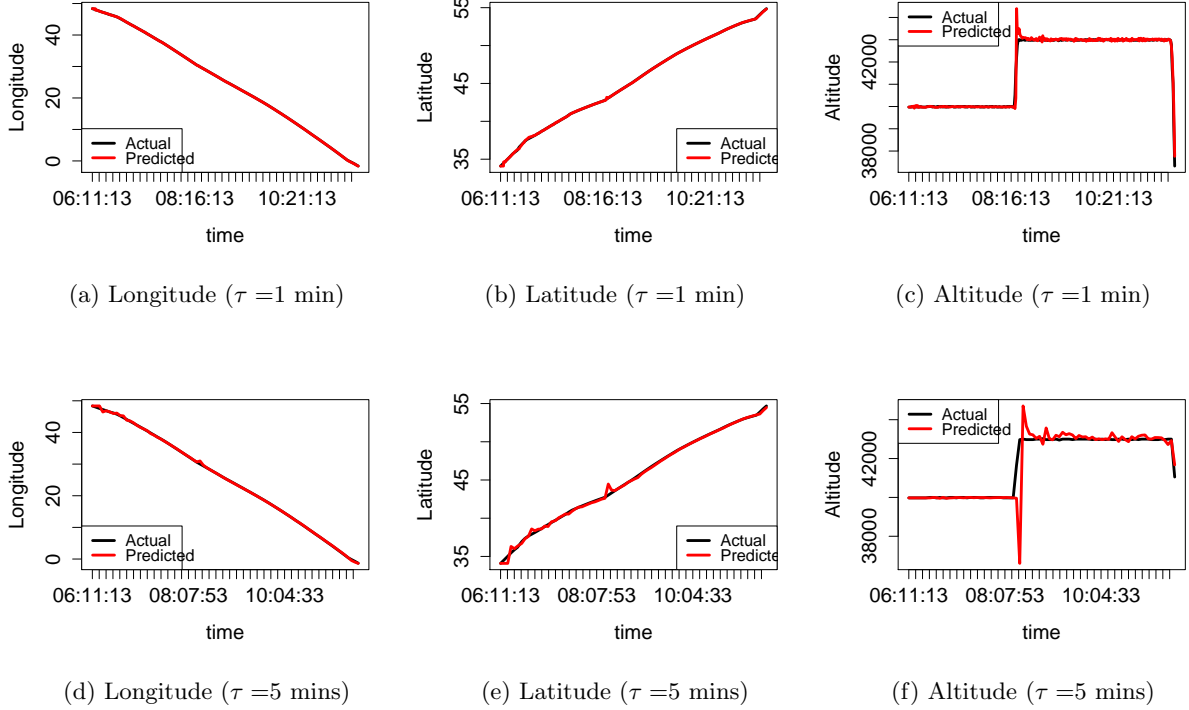


Figure 4: Regression prediction

specifically, we learned the components from two historical journeys of the air route between Doha and London Heathrow airport of aircraft 06A0BC where the data is collected at an interval of every five minutes. As expected, the prediction performs better than the regression approach for longitude and latitude. For altitude, it performs less well only at the beginning of the trajectory.

B. Analysing anomalies with respect to prediction

Once the trajectories are predicted, anomalies can be analysed where the measured metrics are far away from the predicted ones.

For example, when the distances between the predicted and the measured locations are larger than a threshold, it could trigger further investigations. Sudden changes of the altitude, e.g., could be “normal” operations of the pilots during a flight, however, when combining with sudden changes of longitude/latitude from the predicted trajectory, it can mark them as alarming.

To analyse anomalies would typically require full data records of locations for every interval. Depending on the update frequency, however, it may exceed the bandwidth of existing satellite communication technology. Therefore, we experiment some candidate LSH functions to see whether losing precision does not necessarily lead to a loss of accuracy.

C. Rounding hash for verification

A general approach to satisfy the defined location-sensitive verification requirement is to “hash” both actual and predicted aircraft locations in such a way that close locations are more likely to be hashed to the same value than far-away locations are. Specifically in the aircraft motion scenario, we are interested in detecting abnormal motions in longitude, latitude and altitude so that by comparing to a threshold d_{lon} (resp. d_{lat} , d_{alt}) we can determine whether an actual longitude (resp. latitude, altitude) is indeed astray from a predicted value. Hence an actual longitude value over d_{lon} distance away from the predicted value is defined as *abnormal* motion assuming the predicted locations are accurate and reflect the normal motion as

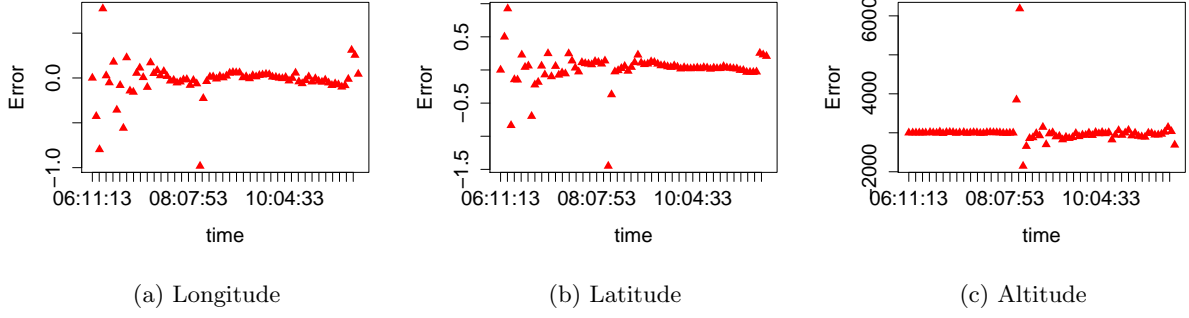


Figure 5: Error of prediction ($\tau = 5$ mins)

shown in Section A. Using a hashing function, we consider any pair of the actual and predicted locations at single timestamps that are hashed to the same value an *identified normal motion pair*, otherwise an *identified abnormal motion pair*. The hope is that most of the identified abnormal motion pairs are truly far-away locations (i.e., the actual abnormal motion), and only a small fraction of close locations could be misidentified as abnormal.

We use the rounding approach to create the hashes of the actual and predicted longitude, latitude and altitude respectively with different levels of rounding accuracy. More specifically, if we focus on the altitude, then $|Altitude_t - Altitude'_t| \leq d_{alt}$ indicates that the actual value $Altitude_t$ and the predicted value $Altitude'_t$ are close. Then a rounding function $h(\cdot)$ with certain accuracy (e.g., 100, 1000) rounds of $Altitude_t$ and $Altitude'_t$ to certain values $h(Altitude_t)$ and $h(Altitude'_t)$. The chosen accuracy of rounding determines the probabilities P_1 and P_2 as follows:

- if $|Altitude_t - Altitude'_t| \leq d_{alt}$, then $h(Altitude_t) = h(Altitude'_t)$ with probability at least $P_1 = \frac{\# \text{ of correctly identified normal pairs}}{\# \text{ of actual normal pairs}}$;
- if $|Altitude_t - Altitude'_t| > cd_{alt}$ where $c = 1$, then $h(Altitude_t) = h(Altitude'_t)$ with the probability at most $P_2 = \frac{\# \text{ of mistakenly identified normal pairs}}{\# \text{ of identified normal pairs}}$,

where $P_1 > P_2$.

We used the trajectory data of aircraft ‘06A0BC’ from Doha to Edinburgh airport to evaluate the performance of the rounding approach. We set the threshold d_{lon} , d_{lat} and d_{alt} as 0.1, 0.1 and 400 respectively, and rounded of each pair with different levels of rounding accuracy, (1, 5, 10) for longitude and latitude and (400, 600, 800, 1000, 1200, 1400, 1600) for altitude. The changes of LSH probabilities P_1 and P_2 are shown in Figure 6, indicating that the LSH condition $P_1 > P_2$ always holds for any rounding (accuracy) parameter. For longitude and latitude, the P_1 is always 0.99 or one and P_2 is always zero. Therefore, the rounding parameter should be set as 10 to obtain minimum communication cost.

To choose the rounding parameter of the best performance in identifying abnormal motion pairs, we look at the recall and precision scores when different levels of rounding accuracy are implemented. As shown in Figure 7a, the recall score for identifying abnormal motion in altitude increases as the rounding parameter changing from 400 to 1,000. Considering precision and recall together with a preference for recall, the rounding approach has the best overall performance when the accuracy is 500, where the recall score is 1 and precision score is 0.8.

D. Update the scalability and liveness requirements

Using the proposed LSH scheme to replace the original full representation of locations, one can get reasonably good precision/recall performance. Now the question is, how much can it enhance the scalability and liveness requirements. Since there are room to improve accuracy by choosing different levels of intervals on the tracking requirement, here we aim to have a rough estimation rather than a precise estimation. Based on the size of flight radar data, earlier we have estimated that it costs at least 514,784 Kbps to transmit the locations of all 21K aircraft in operation through satellites, if the interval is every 15 minutes. Now, if one

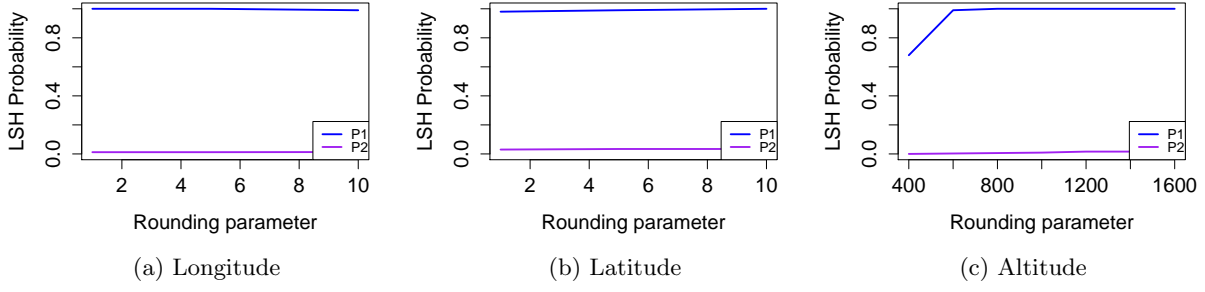
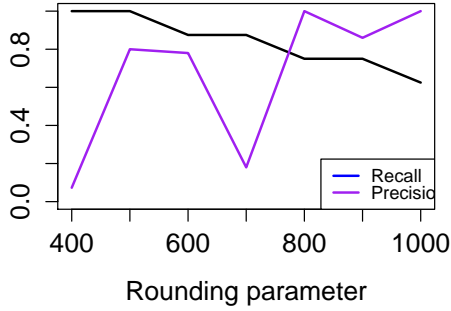


Figure 6: LSH probabilities



(a) Recall and precision for identifying abnormal motion

Rounding Parameter (longitude/latitude)	1-9	10
Bandwidth Saving	5B	6B

Rounding Parameter (altitude)	400-900	1000
Bandwidth Saving	2B	3B

(b) Bandwidth Saving

Figure 7: Effect of applying an LSH encoding

reduces the monitoring interval to every 1 minutes, it would increase the bandwidth requirement to 7,540 Mbps. Further, to live update at 50 ms (before interpolation), the bandwidth requirement could be increased further to 150,800 Mbps. This had not yet include the other data from Flight Data Recorders, such as cockpit voices, passenger voices, etc. On the other hand, the use of Rounding function, the bandwidth requirements can be cut by at least half, while computing the Rounding function has the time/spatial complexity of $\mathcal{O}(1)$.

With the compression of LSH, it is estimated a saving from 19B per data point (i.e., longitude and latitude contribute 7B respectively, and altitude contributes 5B according to the collected data) to 4B, thus a 4.75x saving of the bandwidth in these scenarios. Table 7b details the bandwidth saving while varying the rounding accuracy parameter. To achieve a recall score of 100%, we selected 10 (and 500) as the rounding parameters for longitude/latitude (and respectively for altitude). Therefore, the compressed bandwidth per data point is reduced to 4 bytes.

As shown in the results, there is indeed a correlation between these five requirements. In particular, the change of interval τ affects all the five requirements. When appropriate LSH function is applied as we proposed, however, the verification requirement could be satisfied on smaller intervals, when the other four requirements remain mostly the same.

E. Threats to Validity

CONSTRUCT VALIDITY Amongst the 21K aircraft, some were only recorded for one journey during the last three weeks. For such flights, even though decomposition analysis is believed to be more effective to reduce errors by periodical data, the alternative approach to prediction by the previous location would be more applicable. Despite that, the accuracy shown in the prediction of both methods can tolerate coarser granularities introduced by rounding hashes. Although both regression and decomposition of time series

methods are widely used in the prediction of trajectories for forecasting, we are aware of other approaches, such as online learning, which may have better accuracy in the prediction. However, the aim here is only to show that statistical methods can be applied to improve on the prediction and verification.

Our internal tests show that the precision score is very low had we implemented the LSH function by converting location to binary data and comparing the Hamming distance between the converted actual and predicted locations. The very fact that rounding function of our intuitive choice is better than Hamming distance function suggests that choosing an LSH function could lead to further improvement. The experimental results presented here do not compare the effectiveness of all the commonly used LSH functions, which is left as our future work.

Although the ICAO requirement is for $\tau = 15$ minutes, we choose to perform the experiments on $\tau = 5$ minutes only. This is to have enough data points for performing the experiment on datasets collect over 3 weeks time. If we had more data collected, it would be possible to compare it with $\tau = 15$ minutes. Had we compared with the longer interval, prediction results for longitude and latitude should remain the level of accuracy, but for altitude, the accuracy could drop, which would be future work for us. Overall, the bandwidth savings would be even higher had we compared with the longer intervals.

EXTERNAL VALIDITY The dataset was obtained by FlightRadar directly from globally distributed network of ADS-B receivers monitoring the flying-by aircraft. Relying solely on radar data source could be a risk to the validity of our findings, nonetheless the availability of data on densely populated lands is constantly improving. Due to data sensitivity, we cannot release the entire datasets as is without permission. After careful anonymisation through adaptive sharing¹⁶ and cloud-based privacy protection,¹⁷ we will be able to provide the fully captured tracks of aircraft of the period of the past three weeks to peer researchers.

V. Related work

For global civilian aircraft in motion, we propose to monitor the phenomena at the boundary of cyber- and physical-spaces into data accurate enough for diagnosing the departures from predicted trajectories. It applies the approach to monitoring and diagnosing software requirements¹⁸ generally to systems in physical domains and to problems at global scale.

To the best of our knowledge of the literature, there has not been any similar work.

In this section, we will examine the related work in the areas of aviation safety and security requirements, software verification and validation requirements in aerospace, software uncertainty and accuracy requirements, and the verification of large volume of transactional data.

AVIATION SAFETY AND SECURITY REQUIREMENTS Nuseibeh et al¹⁹ proposed that requirements techniques such as trust assumptions²⁰ and structured argumentation²¹ can be combined to enhance the safety and security of air traffic control (ATC) systems. It emphasised that “one reason that an analyst may fail to construct a convincing argument is that there is not enough information available to justify some claim.” Indeed, without sufficiently rich and verified air traffic data, it is hard to instantiate decisions made to implement security requirements. Furthermore, Lockerbie et al²² and Maiden et al²³ combined the i* method with satisfaction argumentation to analyse dependable ATC systems such as Departure Management (DMAN), from the ground, what could go wrong in the design processes. Following an ontology on security requirements to analyse and support lifelong evolving critical systems such as Arrival management (AMAN) and DMAN, Bergmann et al²⁴ used event-condition-action rule-based change management tools; Yu et al²⁵ proposed an automated approach to incorporating risk assessment into security requirements analysis. Recently, Gramatica et al²⁶ proposed to extend the strategic dependency analysis to analyse the economic drives for airline stakeholders to enhance aviation security.

Before the incident of MH370, most commonly studied aviation safety requirement for ATC is about collision avoidance or prevention, which focuses mostly on the separation of aircraft at a safe distance, either in flight or on runways for AMAN and DMAN. They apply to a few aircraft physically close to each other. In this work, however, we focus on the new challenge in tracking and verifying all aircraft in motion using live blackboxes, which could greatly improve the predictability of the global civilian aviation industry and thus could become one of economic concerns for airline stakeholders.

AEROSPACE SOFTWARE VERIFICATION AND VALIDATION REQUIREMENTS For validating the design of unmanned aircraft systems, it is important to characterise message latency, and existing work has focused

on constructing a distributed simulation environment.²⁷ When the aircraft are manned, simulation cannot replace real-time monitoring of aircraft at the global scale.

Formal statistical model checking has been applied in verifying the correctness of aerospace software systems from behaviour aspects such as scheduling in a stochastic environment.²⁸ Our proposed verification requirements are currently limited to verifying the recorded flight data against their prediction, without looking into the implementation of aviation software systems.

UNCERTAINTY AND ACCURACY REQUIREMENTS The combination of security requirements and their satisfaction argumentation arises from the perspective of insider stakeholders, i.e., designer of the ATC systems. The uncertainty of aviation security manifested by MH370, however, may not be fully satisfied resorting to existing technology due to uncertainty.

The tracking requirements listed here are of high level initially, however, used at large scale with limited bandwidth in satellite communications, a trade-off between information accuracy and precision²⁹ is needed to deal with uncertainties. Letier et al²⁹ suggested that information accuracy is what matters to decision making for software architectures. Likewise, it is considered more important in judging whether the aircraft is on track or not, instead of the exact location with absolutely high precision. For scalability and liveness, if possible, one could and perhaps should sacrifice some unnecessary precision if the judgement is still accurate, while saving the required bandwidth.

VERIFICATION OF LARGE-SCALE TRANSACTIONAL DATA Verifying large quantities of transactions can be achieved for crypto-currencies,³⁰ where the correctness of every transaction can be verified by using the associated hashing values of the “proof of work”. Counter-intuitively, verifying digital values using crypto-hashing functions such as SHA-256 may not be ideal for the continuous data values collected from cyber-physical phenomena, because a slight perturbation of input could cause different decisions. Here, we proposed a slightly different verification requirement based on the locality-sensitive hashing concepts. The current location of an aircraft is regarded abnormal only when its distance to the location predicted by the trajectory is larger than a threshold, with a higher probability than otherwise. In this way, large amount of locations of all aircraft in motion could be efficiently communicated by limited satellite bandwidth because they are mathematically equivalent to verify whether the locations of aircraft are highly likely to be abnormal by comparing the LSH values.

VI. Conclusion and future work

In this paper, we have proposed a handful of ‘live blackbox’ requirements for tracking and verifying the locations of aircraft in motion. By varying the sampling intervals, we have used FlightRadar24 data per minute to assess the scalability of satellite communication, i.e. whether it could track aircraft over remote oceans at an interval below the 15 minutes threshold newly recommended by ICAO. The results have shown that a simple locality-sensitive hashing (LSH) technique (rounding) is already effective in compressing the communicated data stream while still verifying the predicted trajectories at runtime.

Since advanced LSH functions have been applied to compress multi-dimensional data,³¹ we will investigate whether our simple scheme could also be extended to non-location flight data including, e.g., voice recordings in cockpits, improving the granularity of verification further.

Acknowledgement

The authors thank Michael A. Jackson for comments on the early draft of the paper, and FlightRadar24 for offering the datasets on the aircraft in motion. The work is supported by ERC Advanced Grant 291652 on Adaptive Security And Privacy and NPRP 05-079-1-018 from the Qatar National Research Fund (QNRF) on Adaptive Information Security.

References

¹Y. Yu, “The aftermath of the missing flight MH370: what can engineers do?” *Proceedings of the IEEE*, vol. 103, no. 11, pp. 1948–1951, 2015.

²planecrashinfo.com, “Aviation accident statistics.” [Online]. Available: <http://www.planecrashinfo.com/cause.htm>

- ³“Missing Malaysia plane: What we know,” *BBC*, Jan. 2015. [Online]. Available: <http://www.bbc.co.uk/news/world-asia-26503141>
- ⁴V. Walt, “The egyptair black boxes have been detected. but black boxes are really outdated.” [Online]. Available: <http://time.com/4354099/egyptair-flight-804-black-boxes>
- ⁵“MH370 search: Does debris solve the mystery?” [Online]. Available: <http://www.bbc.co.uk/news/world-asia-33713885>
- ⁶“MH370: Reunion wing debris ‘certainly’ from missing flight.” [Online]. Available: <http://www.bbc.com/news/world-asia-34145127>
- ⁷“Washed-up debris from same type of plane as missing MH370, official says.” [Online]. Available: <http://www.cbc.ca/news/world/mh370-debris-found-1.3472949>
- ⁸C. Irving, “Expert: Replace black boxes with new technology.” [Online]. Available: <http://am.blogs.cnn.com/2009/06/09/expert-replace-black-boxes-with-new-technology/>
- ⁹D. van’t Kaar and R. M. Schnitker, “Online black box-legal aspects with regard to real time streaming of critical aircraft flight data/online black box: Rechtliche aspekte der direktübertragung unfallrelevanter flugdaten/online black box: Aspects juridiques relatifs a la transmission directe de l’information pertinente en cas de l’accident aerien,” *ZLW*, vol. 59, p. 513, 2010.
- ¹⁰Y. Yu, “If we’d used the cloud, we might know where MH370 is now.” [Online]. Available: <http://theconversation.com/if-wed-used-the-cloud-we-might-know-where-mh370-is-now-24542>
- ¹¹D. Damon, “How Cloud Computing Could Help Locate the Missing MH370 Plane, an Interview with Dr Yijun Yu from The Open University,” Apr. 2014. [Online]. Available: <http://www.bbc.co.uk/programmes/p01wl74m>
- ¹²S. Govindasamy and T. Leong, “Malaysia airlines upgrades aircraft tracking, a year after MH370.” [Online]. Available: <http://www.reuters.com/article/us-malaysia-airlines-minister-idUSKBN0M20X520150306>
- ¹³M. Datar, N. Immerlica, P. Indyk, and V. S. Mirrokni, “Locality-sensitive hashing scheme based on p-stable distributions,” in *Proceedings of the Twentieth Annual Symposium on Computational Geometry*, ser. SCG ’04, 2004, pp. 253–262.
- ¹⁴A. Aijaz, M. Dohler, A. H. Aghvami, V. Friderikos, and M. Frodigh, “Realizing the tactile internet: Haptic communications over next generation 5g cellular networks,” *CoRR*, vol. abs/1510.02826, 2015.
- ¹⁵M. G. Kendall, *Time-Series*, 2nd ed. Charles Griffin, 1976.
- ¹⁶M. Yang, Y. Yu, A. K. Bandara, and B. Nuseibeh, “Adaptive sharing for online social networks: A trade-off between privacy risk and social benefit,” in *Trust, Security and Privacy in Computing and Communications (TrustCom), 2014 IEEE 13th International Conference on*, Sept 2014, pp. 45–52.
- ¹⁷M. Barhamgi, A. K. Bandara, Y. Yu, K. Belhajjame, and B. Nuseibeh, “Protecting privacy in the cloud: Current practices, future directions,” *IEEE Computer*, vol. 49, no. 2, pp. 68–72, 2016.
- ¹⁸Y. Wang, S. A. McIlraith, Y. Yu, and J. Mylopoulos, “Monitoring and diagnosing software requirements,” *Autom. Softw. Eng.*, vol. 16, no. 1, pp. 3–35, 2009.
- ¹⁹B. Nuseibeh, C. B. Haley, and C. Foster, “Securing the skies: In requirements we trust,” *IEEE Computer*, vol. 42, no. 9, pp. 64–72, 2009.
- ²⁰B. Haley, C. Laney, D. Moffett, and B. Nuseibeh, “Using trust assumptions with security requirements,” *Requir. Eng.*, vol. 11, no. 2, pp. 138–151, Feb. 2006.
- ²¹C. B. Haley, J. D. Moffett, R. Laney, and B. Nuseibeh, “A framework for security requirements engineering,” in *Proceedings of the 2006 International Workshop on Software Engineering for Secure Systems*, ser. SESS ’06, 2006, pp. 35–42.
- ²²J. Lockerbie, N. A. M. Maiden, J. Engmann, D. Randall, S. Jones, and D. Bush, “Exploring the impact of software requirements on system-wide goals: a method using satisfaction arguments and i* goal modelling,” *Requir. Eng.*, vol. 17, no. 3, pp. 227–254, 2012.
- ²³N. A. M. Maiden, J. Lockerbie, D. Randall, S. Jones, and D. Bush, “Using satisfaction arguments to enhance i* modelling of an air traffic management system,” in *15th IEEE International Requirements Engineering Conference, RE 2007, October 15-19th, 2007, New Delhi, India*. IEEE Computer Society, 2007, pp. 49–52.
- ²⁴G. Bergmann, F. Massacci, F. Paci, T. T. Tun, D. Varró, and Y. Yu, “Secmer: A tool to gain control of security requirements evolution,” in *Towards a Service-Based Internet - 4th European Conference, ServiceWave 2011, Poznan, Poland, October 26-28, 2011. Proceedings*, ser. Lecture Notes in Computer Science, W. Abramowicz, I. M. Llorente, M. Surridge, A. Zisman, and J. Vayssi re, Eds., vol. 6994. Springer, 2011, pp. 321–322.
- ²⁵Y. Yu, V. N. L. Franqueira, T. T. Tun, R. Wieringa, and B. Nuseibeh, “Automated analysis of security requirements through risk-based argumentation,” *Journal of Systems and Software*, vol. 106, pp. 102–116, 2015.
- ²⁶M. de Gramatica, F. Massacci, W. Shim, A. Tedeschi, and J. Williams, “IT interdependence and the economic fairness of cybersecurity regulations for civil aviation,” *IEEE Security & Privacy*, vol. 13, no. 5, pp. 52–61, 2015.
- ²⁷J. Murphy, “Message latency characterization of a distributed, live, virtual, constructive simulation environment,” *SOFTWARE CHALLENGES IN AEROSPACE, SCITECH, AIAA*, 2015.
- ²⁸J. Hansen, L. Wraga, S. Chaki, D. de Niz, and M. Klein, “Verification of real-time systems using statistical model checking,” *SOFTWARE CHALLENGES IN AEROSPACE, SCITECH, AIAA*, p. 11, 2015.
- ²⁹E. Letier, D. Stefan, and E. T. Barr, “Uncertainty, risk, and information value in software requirements and architecture,” in *Proceedings of the 36th International Conference on Software Engineering*, ser. ICSE 2014, 2014, pp. 883–894.
- ³⁰G. G. Dagher, B. B n z, J. Bonneau, J. Clark, and D. Boneh, “Provisions: Privacy-preserving proofs of solvency for bitcoin exchanges,” in *Proceedings of the 22Nd ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS ’15, 2015, pp. 720–731.
- ³¹Y. Liu, J. Cui, Z. Huang, H. Li, and H. T. Shen, “SK-LSH: an efficient index structure for approximate nearest neighbor search,” *PVLDB*, vol. 7, no. 9, pp. 745–756, 2014.